

## Una herramienta permite clasificar opiniones políticas en Twitter

**La Universidad de Salamanca ha desarrollado un clasificador de tuits en tiempo real que funciona mediante aprendizaje automático y en la nube**

Las redes sociales se han convertido en un espacio primordial para el debate público y entre ellas destaca especialmente Twitter en el ámbito de la política. Algunas herramientas informáticas ya analizan tendencias y opiniones. La Universidad de Salamanca ha desarrollado AutoCop, un clasificador de mensajes que permite conocer en tiempo real las valoraciones positivas y negativas de los usuarios.

“Es un clasificador de tuits políticos en tiempo real que permite distinguir entre sentimientos positivos, neutros y negativos. Lo hemos desarrollado a partir de la técnica de análisis de sentimientos supervisado, que utiliza algoritmos de aprendizaje automático basado en inteligencia artificial”, explica a DiCYT Carlos Arcila Calderón, investigador del Departamento de Sociología y Comunicación de la Universidad de Salamanca.

En una primera fase, los investigadores adaptaron al castellano un código en el lenguaje de programación Python que ya existía en inglés. Para ello, entrenaron al sistema mediante un corpus que recogía ejemplos de tuits políticos en castellano clasificados de forma manual.

Ahora han dado un paso más al subir el clasificador a la nube con la plataforma de código abierto Apache Spark, un desarrollo que no existía en acceso libre en ningún idioma. De esta forma, podrán analizar una gran cantidad de datos desde cualquier parte de forma eficiente y segura.

El nuevo prototipo de AutoCop permitirá abordar problemas de análisis masivo de datos, es decir, lo que se conoce como Big Data. Como está alojado en la nube, supera las limitaciones del ordenador personal para poder procesar, gestionar y analizar en tiempo real los datos de Twitter, ya que se puede hacer de forma distribuida, entre muchos ordenadores, y de forma más rápida.

“No solo es más rápido, sino que es más estable porque no depende de las dificultades que podamos tener en casa o en la oficina, como puede ser que se vaya la luz o se caiga el ordenador, en la nube podemos distribuir y ejecutar todo el código con mucha estabilidad y velocidad”, comenta Arcila.

Para desarrollar esta segunda fase del proyecto, los investigadores del [Observatorio de Contenidos Audiovisuales](#) de la Universidad de Salamanca han participado en la convocatoria Prueba de Concepto de la Fundación General de la Universidad de Salamanca, dentro del programa TCUE de la Junta de Castilla y León cofinanciado con fondos FEDER.

*“Esta actuación se encuadra en el Plan TCUE 2015-2017, y ha sido seleccionada en el marco de un programa operativo cofinanciado por el Fondo Europeo de Desarrollo Regional (FEDER) y la Junta de Castilla y León”*



FONDO EUROPEO DE  
DESARROLLO  
REGIONAL



UNIÓN EUROPEA

### Gráficos en tiempo real

“Una de las principales ventajas de AutoCop, tanto en su versión local como distribuida, es que podemos visualizar los sentimientos en tiempo real, es decir, contamos con gráficos que nos indican qué sentimiento existe ahora mismo hacia un determinado partido político o hacia una determinada idea”, destaca Arcila.

Dadas las características de este clasificador basado en aprendizaje automático, los creadores de AutoCop consideran que puede ser útil tanto para la investigación científica como para su transferencia tecnológica al mundo del periodismo, la consultoría política y los propios partidos, e incluso a otros sectores.

“La aplicación del instrumento de software y de ese *know how* que existe se puede dar en distintos sectores, en comunicación política, responsabilidad social corporativa, incluso marketing y salud pública”, opina Félix Ortega Mohedano, otro de los investigadores del proyecto. “El diálogo de lo que sucede hoy en día en la sociedad se da en los espacios de internet, en particular en Twitter y en otros espacios tenemos que investigar con metodologías científicas y de Big Data como AutoCop”, agrega.

A través de este proyecto, los conocimientos en materias como la política, la comunicación y la sociología se dan la mano con las tecnologías más avanzadas y la inteligencia artificial para conocer mejor la realidad.

*“Esta actuación se encuadra en el Plan TCUE 2015-2017, y ha sido seleccionada en el marco de un programa operativo cofinanciado por el Fondo Europeo de Desarrollo Regional (FEDER) y la Junta de Castilla y León”*



## A tool to classify political opinions on Twitter

**The University of Salamanca has developed a real-time tweet classifier that works in the cloud through an automatic learning process**

Social networks have become a prime space for public debate and among them Twitter especially stands out in the field of politics. Some computer tools already analyze trends and opinions. The University of Salamanca has developed a messages classifier that allows to know in real time the positive and negative evaluations of the users.

It is a real time political tweets classifier that allows distinguishing between positive, neutral and negative feelings. We have developed it from the supervised feelings analysis technique, which uses machine learning algorithms, based AI", explains to DiCYT Carlos Arcila Calderón, researcher at the Department of Sociology and Communication.

In an initial phase, researchers adapted from English to Spanish an existing Python code.

To do this, they trained the system through a corpus that collected examples of political tweets in Spanish which have been manually classified.

Now they have uploaded it to the cloud with the Apache Spark open source platform, a development that was not free access in any language. This will allow them to analyze efficiently and safely a large amount of data from anywhere.

The new AutoCop prototype will address itself massive data analysis processes, also known as Big Data. As it is hosted in the cloud, it overcomes personal computers limitations as it is able to process, manage and analyze data coming from Twitter, very fast, in real time and in a distributed way.

"Not only it's faster, but it is more stable because it does not depend on the difficulties we may have at home or in the office, such as the power going out or the computer crashing, in the cloud we can distribute and execute the analysis with a lot of stability and speed," says Arcila.

To develop this second phase of the project, researchers of the Observatory of Audiovisual Content from the University of Salamanca have taken part in the Proof of Concept call of the General Foundation of the University of Salamanca, within the TCUE program of the Junta de Castilla y León co-financed with FEDER funds.

**Real time graphics**  
"One of the main advantages of AutoCop, both in its local and distributed version, is that we can visualize "Tweeter real time feelings", that is, we have graphics that indicate what feeling exists in a certain moment towards a certain political party or towards a certain idea ", highlights Arcila.

Given the characteristics of this machine learning classifier, the creators of AutoCop consider that it can be applied both for scientific research and for being transfer to other sectors such as journalism, political consulting and the parties themselves.

*"Esta actuación se encuadra en el Plan TCUE 2015-2017, y ha sido seleccionada en el marco de un programa operativo cofinanciado por el Fondo Europeo de Desarrollo Regional (FEDER) y la Junta de Castilla y León"*



FONDO EUROPEO DE  
DESARROLLO  
REGIONAL



UNIÓN EUROPEA

"Our software and know-how may be useful in different sectors, for political communication, corporate social responsibility, including marketing and public health", Félix Ortega Mohedano, one of the project's researchers. "Twitter may be considered the arena for nowadays public dialogue. So that, AutoCop allows us to study what's happening in social media through scientific methodologies and Big or 'medium-sized data', he adds.

Through this project, knowledge in subjects such as politics, communication and sociology go hand in hand with the most advanced technologies and artificial intelligence to better understand reality.

*"Esta actuación se encuadra en el Plan TCUE 2015-2017, y ha sido seleccionada en el marco de un programa operativo cofinanciado por el Fondo Europeo de Desarrollo Regional (FEDER) y la Junta de Castilla y León"*

